Perth, Australia. 1–4 February 1999

# LOW-COST IMPLEMENTATION OF A SPEECH COMPRESSION SYSTEM

**D. Gibson**

*Institute of Integrated Information Systems, University of Leeds, Leeds, LS2 9JT, UK*
D.Gibson@caves.org.uk

**ABSTRACT**: A method is described for compressing analogue speech waveforms to reduce the peak/mean ratio without adversely affecting the speech intelligibility or acceptability. The software is relatively simple and is suitable for implementation in a low-cost RISC microcontroller. Methods of testing the resultant intelligibility using a diagnostic rhyme test or phonetically-balanced world list are described.

## 1. INTRODUCTION

This paper describes a method of compressing analogue speech that reduces the peak/mean ratio of the waveform without adversely affecting the speech intelligibility or acceptability. In particular it will describe a low-cost, low-complexity implementation that brings the technique within reach of radio amateurs, electronics enthusiasts and other experimenters. Although DSP techniques are used, the compression technique is essentially that of an analogue speech processor; no digital data compression is involved.

The technique described here was developed over twenty years ago at the University College of Swansea by Dr Louis Thomas [1]. It has been reported on several occasions in the amateur electronics press [2, 3, 4] and the British Technology Group (BTG) has provided development funding and patent provision [5, 6].

The speech compressor takes the form of a feedforward mean-loudness controller and waveshape compressor, and is known as ***Similar*** (for **sim**ultaneous, near-**i**nstantaneous and **t**ime-**a**veraged **r**esponse). Some intriguing claims are made for this system and, apart from commercial uses, it would provide interesting experimentation for radio amateurs and electronics enthusiasts. For this, it would be desirable to produce an implementation that did not require the use of a complex DSP and its supporting equipment for programming and simulation.

## 2. SPEECH PROCESSING METHODS

A technique used in the past for increasing speech intelligibility in a radio channel was to clip the signal with a pair of diodes. By limiting the level of the larger signals the quieter signals could be resolved, but at the expense of increased background noise and severe harmonic and intermodulation distortion. Back-to-back diodes across an op-amp give a 'softer' clip than a simple clamp to ground, and a so-called 'sideband' or 'r.f.' clipping method, e.g. [7], shifts some of the harmonic distortion to beyond the audio band, but this is still far from adequate.

Other methods of controlling the signal level using ALC compression and by the transmission of tones have been tried, with more success, but there are still limitations.

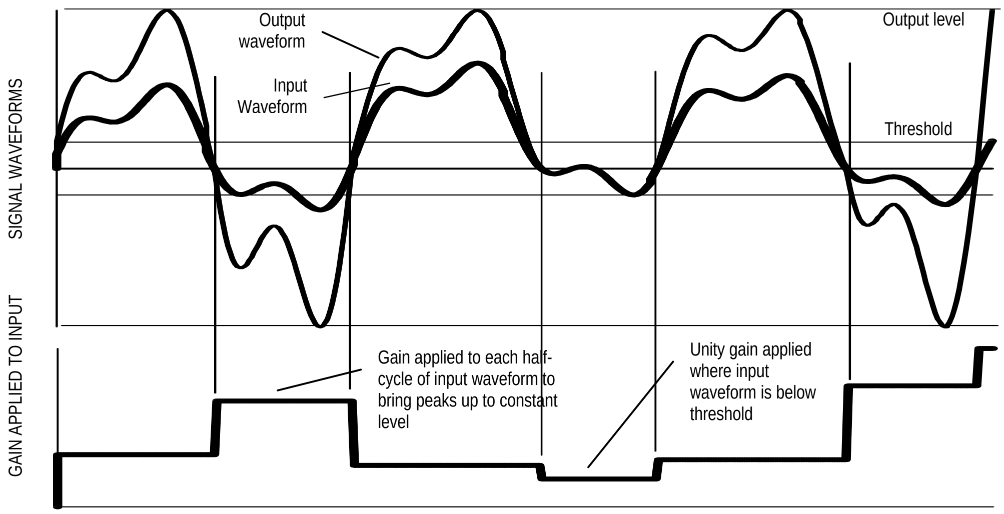
Thomas's technique is essentially that of a very fast-acting compressor but, to avoid intermodulation distortion, gain changes are applied only on the zero-crossings of the signal. This requires the waveform to be sampled and stored for a brief period to allow a feed-forward gain control to act.

## 3. THE 'SIMITAR' SYSTEM

The principle of Simitar is illustrated in ***Figure 1*** below. Each half cycle of the input waveform is boosted to bring its peak up to a constant level. Signals that do not exceed a threshold are ignored in order to avoid amplifying the background noise. Because this compression acts so fast, but without degrading the signal, it suggests some interesting uses. The system was initially claimed to 'punch up' the average loudness of a communications system by as much as 20dB without exceeding peak modulation levels. It was also claimed to increase the mean power level by 6dB and by 15dB on speech consonants.

As well as uses in radio communications, it is possible that the system has studio and broadcast uses. Because the technique boosts consonant levels, which are the quieter parts of a speech waveform, it may be suitable for public address systems – Thomas has reported encouraging results of tests with people with impaired hearing [4]. Although mainly intended for use at the transmitting end of a communications link, its use as a deaf-aid is one example where the device can be placed at the receiver to improve intelligibility in a noisy channel, where the transmitted signal cannot be accessed for processing.

***Figure 1 – a waveshape compressor acting on each half-cycle of the input waveform***




# 4. IMPLEMENTATION

In 1976, the technique was difficult to implement, requiring a large number of discrete components. In 1993 Thomas reported that the process had been implemented using a *TMS320C17* DSP with external PCM codecs on a 110 × 116mm PCB. Today, the entire process could probably be implemented in a single device. One possible contender would be the *RSC164* from Sensory Circuits' *Interactive Speech*™ series of products. The 68-pin PLCC package houses a 'fully functional' 8-bit microcontroller with on-chip RAM and ROM, on-chip DAC/ADC and architecture tailored for speech applications, including recognition, synthesis and voice record/playback.

The possibilities for this method of speech processing are so intriguing that it would be useful to open the door to amateur experimenters; and a professional surface-mount IC is, perhaps, a bit restricting. But the basic algorithm for the waveshape compressor is extremely simple, containing only a handful of instructions [8], which can be implemented in only a few hundred machine cycles. The main difficulty is in the division operation needed to adjust the gain, but this can be avoided by programming an external multiplying DAC, as shown in ***Figure 2*** below.

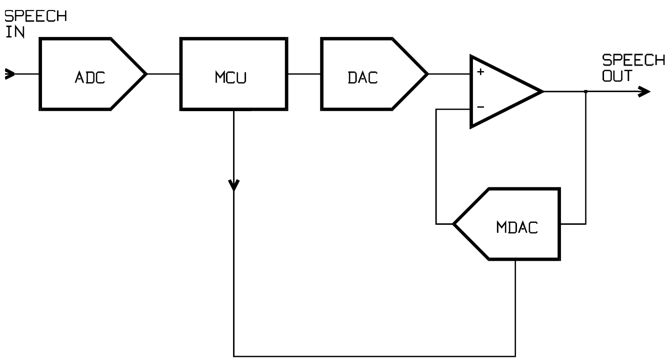



***Figure 2 – a low-cost implementation using a RISC micro-controller (MCU) and a multiplying digital-to-analogue converter (MDAC)***

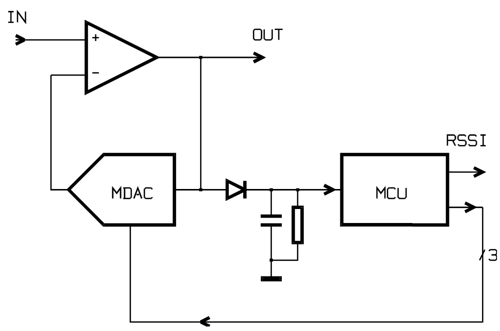



***Figure 3 – the implementation of a conventional ALC/AGC circuit using a mixed-signal RISC microcontroller and MDAC***

It thus becomes possible to build a speech compressor using a simple, low-cost RISC microcontroller, such as one of Arizona Microchip's PIC range. The advantage is that the system is easy to build and easy to program. For commercial use, this multi-IC approach would not be cost-effective; and that is where a single DSP chip, such as mentioned above, would be preferred.

# 5. TESTING, AND FURTHER DEVELOPMENT

### 5.1 Speech Intelligibility

Additional software routines would allow the MCU to report the input and output peak/mean ratio although, with a simple micro-controller, this could be a burden on the processor. To test the claim that the technique increases intelligibility by boosting speech consonants requires a subjective test. There are a number of such tests available including the 'diagnostic rhyme test' (DRT) and the 'phonetically balanced word list' (PB); both of which have ANSI specifications [9]. for a PB speech test a set of words is read out, noted down by the subject, and used to score an 'articulation loss' of consonants (%ALcons). If there were, say, 15 words in error in a list of 100, this would be a 15% loss of articulation. 15% Alcon is generally considered unintelligible although some engineers consider 10% unacceptable. A suitable PB word list appears in [10, 11].

These, and other similar tests, are used to evaluate speech *intelligibility*. This is not quite the same thing as speech *quality* or *acceptability*, which is measured by a separate category of tests, such as Diagnostic Acceptability Measure (DAM), Mean Opinion Score, and so on. These are similar to the radio amateur's SINAD scoring method. The DAM was developed at Dynastat Inc. and more information about speech quality testing can be found on their web site [12].

### 5.2 Simitar Developments

**Waveshape compressor**. The implementation described so far is a 'near instantaneous' method of speech compression, since it acts on each half cycle of the waveform as soon as it has been received. Although this achieves the aims of a) decreased peak/mean ratio and b) increased intelligibility over conventional clipping methods, it does cause speech to have a characteristic and slightly odd timbre. However, a modification to the concept allows the system to act more like a conventional 'automatic level control' (ALC).

**Loudness Controller**. Thomas described how the queue of signal peaks, used to derive the near-instantaneous gain, could instead be applied to a simple first order digital filter. The filter is synchronised to the zero-crossings of the waveform to preserve the characteristics of the system, and has a controllable attack and decay. In addition, the output of the filter can be used to program a gain function rather than setting the gain directly. In this way, functions such as limiting, compression, expansion and noise-gating are possible.

Both these methods – waveshape compressor and loudness controller can be implemented together and the parameters adjusted to suit the application. Further detail, including block diagrams of the system and experimental results can be found in Thomas's articles and patents. It is not necessary to go into detail here.

### 5.3 Conventional AGC/ALC

A similar, but simpler, architecture to that of ***Figure 2*** allows us to implement a conventional AGC or ALC circuit. The ADC, DAC and samples queues are not needed and we can, instead, provide the MCU with just a rectified and filtered audio envelope. ***Figure 3***, above, shows one possible architecture; the MCU's internal ADC is sufficient and, to allow this to operate efficiently, it is driven from the relatively constant level present at the output of the system. This means that the software processing inside the MCU is slightly different than for a 'conventional' AGC consisting of rectifier, filter and variable gain amplifier.

The advantage of using an MCU is that it allows us to simulate AGC and ALC functions similar to those provided by the now obsolete Plessey SL621C AGC Generator [13], which was much used by radio amateurs. This useful chip was intended for monitoring the audio output of an SSB receiver for which it generated a control voltage to feed back to the RF amplifier (SL610 etc.). It operated using five time-constants and a set of analogue comparators, arranged so that it would attenuate noise spikes, but would, nevertheless, track the audio. When the audio disappeared, the AGC setting was retained. Clearly this is rather simpler to manage in software than – now that the IC is obsolete – in hardware. The Simitar implementation of ***Figure 2*** would normally be placed at the transmitter end of a link, whereas the AGC implementation of ***Figure 3*** would probably find better use at the receiver.

# 6. EXPERIMENTAL RESULTS

The original system underwent some years of development and was evaluated by its inventor, Dr. Louis Thomas, for peak/mean ratio; real and subjective SNR and speech degradation. It is not our immediate intention to attempt to repeat this work, but simply to test the feasibility of a novel low-cost implementation.

Initial tests with a PIC microcontroller from Arizona Microchip are being carried out at the moment. It is planned to present some simple results and to give a live demonstration of the performance of the low-cost implementation at DSPCS '99. A 'single chip' implementation using a Sensory Circuits RSC164 is planned for the future.

# 7. CONCLUDING REMARKS

A waveshape compressor and mean loudness controller for analogue speech have been described. The control functions could be implemented in a DSP, and we would expect to do that in a commercial product. But sometimes, analogue functions are best implemented in analogue devices, and so we have described a low-cost easy-to-build solution that can be used for further experimentation, and which may be attractive in some applications.

# 8. ACKNOWLEDGEMENTS

The system described here was invented by Dr Louis Thomas and has been explained at some length in articles published by him. Patents [5, 6] covering aspects of this work have lapsed and the British Technology Group reports that no company is currently licensed to use this technology. Neither the BTG nor Dr Thomas's former colleagues at the University of Wales have been able to contact him recently.

# 9. REFERENCES

1 Thomas, L.D. (1976) *an investigation into a non-linear method of automatic level control for speech channel signals*, MSc thesis, University College of Swansea

2 Dance, Brian. (1985) *Digital speech processing*, Practical Wireless, April 1985, pp20-22, ISSN 0141-0857

3 Thomas, Louis. (1991) *High quality punch for radiocomms*, Electronics World + Wireless World, **97**(1664), July 1991, pp559-560, ISSN 0959-8330

4 Thomas, Louis. (1993) *Putting punch into voice communications*, Electronics World + Wireless World, **99**(1682), Jan. 1993, pp14-20, ISSN 0959-8330

5 Thomas, L.D. (1977) UK Patent Specification 1599401 – Input Signal Level Control for Communications Channels, Filed 22 March 1977

6 Thomas, L.D. (1988) UK Patent GB2201310B – Signal Level Control, Filed 18 Jan 1988

7 Anderson, P. (1979) *Audio processor design*, Wireless World, **85**(1524), Aug. 1979, pp34-9 & 62. ISSN 0043-6062

8 Gibson, David. (1991) *Speech compressor*, Electronics World + Wireless World, **97**(1666), Sept. 1991, p757, ISSN 0959 8330

9 ANSI standard S3.2-1989, *Measuring the intelligibility of speech over communication systems*.

10 Klark Teknik (publisher), Kidderminster, England, (1988), The Audio System Designer, *no ISBN*

11 Gibson, David (1998) *A phonetically-balanced speech intelligibility test*, Cave Radio & Electronics Group Journal **31**, pp14-15, March 98, , ISSN 1361-4800

12 Dynastat, Inc. Austin, Texas, USA. *Subjective testing and evaluation of voice communications systems.* Web site http://www.bga.com/dynastat/index.html

13 Plessey Semiconductors (1988), Linear IC Handbook, publication number P.S. 1973